International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I1P101
Volume 72 Issue 1, 1-10, January 2024

Original Article

A Hybrid Weiner Filter with MR-CNN for Object Detection in Underwater Image Processing

Hemalatha[1], Saravanan[2]

[1,2]*Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, Tamilnadu, India.*

[1]*Corresponding Author : hemalathasundaraj@gmail.com*

Received: 09 November 2023 Revised: 13 December 2023 Accepted: 26 December 2023 Published: 07 January 2024

Abstract - *Underwater image processing poses unique challenges due to the presence of various degradation factors such as color attenuation, backscatter, and noise. Object detection in underwater images is particularly challenging as these factors can affect the visibility and clarity of objects. In this study, we present a unique approach that combines the Weiner filter with a Multi-Resolution Convolutional Neural Network (MR-CNN) for a method for detecting objects underwater image processing to enhance object-detecting ability, and we integrate the Weiner filter with an MR-CNN. The MR-CNN utilizes multiple resolutions to capture and analyze different levels of information in the image. This multi-resolution approach allows for better extraction of features at various scales, enabling the network to detect accurately. The combination of the Weiner filter and the MR-CNN significantly improves the object detection accuracy in underwater images compared to traditional methods. The results highlight the potential for practical applications in underwater research, exploration, and surveillance domains.*

Keywords - *Underwater image processing, Weiner filter, Multi-Resolution Convolutional Neural Network, Object detection, Surveillance domains.*

1. Introduction

Underwater image processing and analysis have gained significance owing to their wide variety of applications; they have acknowledged a lot of interest in the past few years, particularly in marine biology, underwater robotics, and underwater surveillance. However, underwater images pose unique challenges compared to their terrestrial counterparts. Factors such as light attenuation, color distortion, and backscatter cause degradation in image quality, making object detection a complex task [1].

Object detection in underwater images plays a vital role in various domains. It enables the identification and localization of objects of interest, such as marine organisms, underwater structures, or man-made objects [2]. Accurate and robust object detection in underwater environments is essential for tasks like underwater exploration, environmental monitoring, and underwater navigation [3].

Underwater image processing plays a critical role in various applications, including marine research, underwater robotics, and underwater surveillance. However, underwater imageries grieve from challenges such as low visibility, colour distortion, and noise interference, making object detection a challenging task [4]. Traditional image processing techniques, such as the Weiner filter, have been employed to address noise reduction and image restoration [5].

Deep learning, on the opposing hand, is a method of studying, such as CNNs, has demonstrated remarkable success in object detection tasks [6]. The Weiner filter is a classical technique used for noise reduction and image enhancement. It exploits the signal's characteristics and noise's statistics qualities to estimate the original signal and suppress unwanted noise components [7].

By adaptively adjusting the filter parameters, the Weiner filter can effectively restore image details and improve image quality. However, the Weiner filter alone may not fully address the complexities and variations encountered in underwater scenes, where additional challenges such as color distortion and low visibility exist [8].

With recent advancements in deep learning, CNNs have developed effective techniques for object detection. CNNs are learning and extracting high-level landscapes from raw image data, enabling them to detect objects in complex and challenging environments [9][10]. However, directly applying CNNs to underwater images may not provide excellent outcomes due to the unique characteristics and image degradation factors associated with underwater imaging.

This research aims to address these limitations by proposing a novel approach that combines the strengths of a

Hybrid Wiener Filter and a Modified Region-based Convolutional Neural Network (MR-CNN) for improving object detection accuracy in underwater image processing. The scarcity of robust methodologies tailored explicitly for underwater object detection underscores the need for innovative strategies that can mitigate the challenges posed by underwater imaging conditions.

The existing research predominantly focuses on either traditional image enhancement techniques or deep learning-based approaches independently, neglecting the potential synergy achievable through their fusion. Consequently, the research gap lies in the absence of a comprehensive solution that intelligently integrates both image enhancement and deep learning methodologies to enhance object detection accuracy in underwater images.

Following is a brief overview of the remaining sections of the article: Efforts in underwater image processing and identification of objects. Section 3 details the methodology, including the Weiner filter and MR-CNN integration. Section 4 presents the experiment's setup and outcomes. At last, Section 5 summarises the whole article. The contributions and consequences of the suggested technique in Image processing in the underwater environment and object identification are discussed.

2. Related Works

Underwater image processing and object detection are crucial aspects of various fields, including marine biology, oceanography, underwater archaeology, and offshore industries. However, the underwater environment presents unique challenges that significantly hinder traditional image processing and object detection techniques. Underwater images suffer from various degradations, including poor visibility caused by light attenuation, color distortion, scattering, and occlusions due to particles or marine life. These factors collectively contribute to reduced image quality, making it arduous for conventional computer vision algorithms to detect and identify objects in such scenes accurately.

Traditional computer vision techniques, such as edge detection and template matching, have been utilized for object detection in underwater images. However, these methods often struggle to handle the specific challenges present in underwater environments, leading to limited performance and accuracy. In order to get the binary image of the region of interest, an image must first be segmented in accordance with the automated threshold segmentation. By calculating the second instant, an estimate of the target size is made, and the incorrect target is taken out of the equation. Similarly, the precise location of a tiny object submerged in water may be accomplished [11]. The findings of underwater object identification, which is an essential component of the workflow for image processing [12], are very important not only for the upkeep and repairs of underwater structures but also for marine research. To find solutions to challenges such as being sensitive to complicated surroundings and their values, which are developed from drops. When it occurs, it is actually faced with noisy efforts [13]. A method through which designers classify various suggested alterations to object identification by making use of a basic geometric assessment of the scene. As a way of evading the dangerous high-pressure deep-sea environment, underwater autonomous operation is becoming an increasingly vital component. As a consequence of this, it is necessary for there to be an investigation of the ocean floor [14].

In the marine environment, underwater targets are often influenced by interference from other targets and environmental variations; as a result, it is difficult to employ typical target-tracking techniques for the purpose of monitoring underwater objects precisely and reliably [15]. In recent years, the identification of marine fish species has emerged as a significant study field for the preservation of the ocean ecosystem. Finding different types of marine fish on the ocean bottom by human inspection is a laborious and time-consuming process [16]. It is anticipated that the findings of the proposed framework will contribute to the development of subsequent studies that will provide new insights into dolphin behavior, biomechanics, and the effects of the environment on dolphin activity and movements. The work that has been presented here [17] demonstrates the feasibility of tracking marine mammals using CNN object detection.

Generic CNN-based object detectors still have a long way to go before they can successfully identify objects in the underwater environment. This is because the underwater world is so complicated. Low accuracy and recall rates are the consequence of these problems, which include image blurring, texture distortion, colour shift, and size fluctuation [18]. Because of deep learning, the path has been set for improved background subtraction, which will help combat the big obstacles that are present in this area. Additionally, the combination of numerous characteristics results in an enhancement of traditional approaches for background subtraction [19]. Because of the unique qualities of underwater settings, recognising and tracking underwater objects may be a difficult task. When light travels to greater depths in water, the light will diffraction and scatter as it passes through the water medium [20].

The Kalman filter is used with Faster R-CNN and YOLO v2 in the research described in this publication [21], which results in improved detection accuracy. Because Faster R-CNN produces more accurate results, those results are used as observations, whereas those produced by YOLO v2 are used to define state variables. A method of image processing is suggested to be used in the design of a control system for underwater object tracking that is included in this work [22].

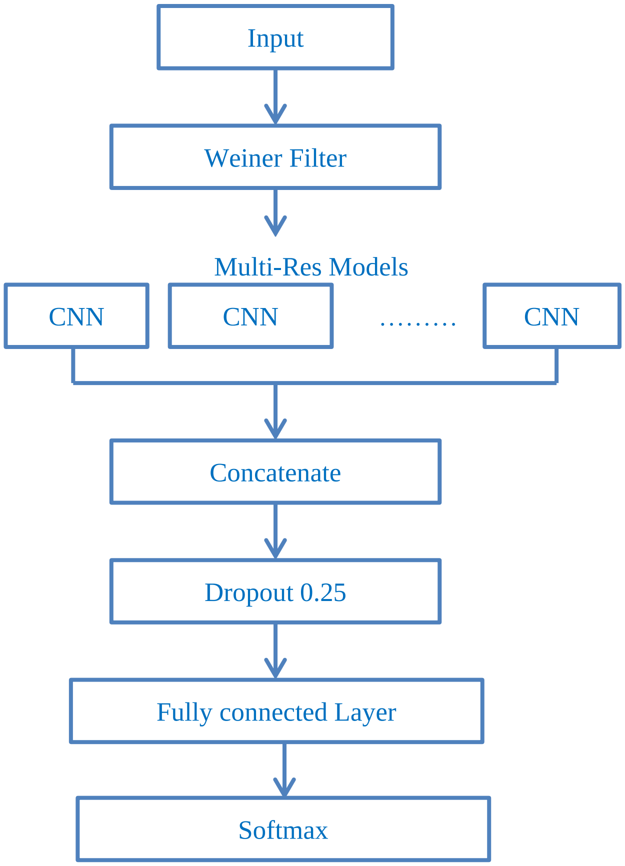

Fig. 1 Proposed model

It is used for the detection of autonomous underwater vehicles at close range as well as the dynamic tracking of these vehicles. Counting objects is an essential duty in aquaculture, and it has been regularly engaged in the estimation of fish populations, estimates of lobster abundance and scallop stocks, and a variety of other estimations [23].

When confronted with varying environmental conditions, fish will swim at varying speeds. Obviously, the rate at which fish swim may be affected not only by their own behaviour and the status of their health but also by the quality of the water in which they swim [24]. The effects of object tracking become increasingly difficult to achieve in underwater films due to the fact that the hazy backdrop, lighting situation, and occlusion significantly impact the movies captured. Therefore, the purpose of this study [25] is to investigate an efficient method for monitoring several items underwater. The primary benefit of this method is its ability to associate objects for online and real-time purposes successfully.

Historically, traditional methods like image enhancement techniques (e.g., filtering and contrast enhancement) have been employed to mitigate the effects of these degradations. However, these techniques often fall short in handling the complex and dynamic nature of underwater scenes, leading to limited success in object detection.

The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), has shown remarkable success in various computer vision tasks. Yet, applying CNNs directly to underwater images encounters challenges due to the lack of annotated data specific to underwater environments and the need for adaptations to accommodate the unique characteristics of underwater imagery.

The intersection of traditional image processing methods and deep learning approaches for underwater object detection remains largely unexplored. Bridging this gap is vital to enhance the accuracy, reliability, and adaptability of object detection systems in underwater settings. Researchers are seeking innovative methodologies that integrate image enhancement techniques with deep learning architectures to address the complexities inherent in underwater image processing and object detection.

3. Proposed Model

The proposed approach involves the following steps: First, the input underwater image is pre-processed using the Weiner filter to reduce noise and enhance relevant information. Next, the pre-processed image is fed into the MR-CNN, which consists of multiple pathways operating at different resolutions. The MR-CNN learns discriminative features from the input image, enabling accurate object detection. Finally, the integrated features are used for object detection through fully connected layers in the network, as shown in Figure 1.

3.1. Wiener Filter

In image processing, it is common to use the Wiener filter to denoise images and restoration tasks. It aims to reduce noise as well as enhance the overall visual clarity of the images by estimating the original, undistorted image from the observed noisy image. The Wiener filter operates in the frequency domain and takes advantage of the statistical properties of the image and noise.

The mathematical equation for the Wiener filter in image processing can be represented as follows:

Let's assume a noisy image, denoted by Y, and want to evaluate the original image, denoted by X, from the noisy observation. In the frequency domain, the Wiener filter operates on the Fast Fourier transformed images. Let Y(f) and X(f) signify A noisy image is transformed using the Fast Fourier transform and the original image, correspondingly. Similarly, let N(f) represent Additive noise transformed using the Fast Fourier transform.The proportion of the power spectra of the clean image to those of the noisy image is the definition of the Wiener filter transfer function, which is indicated by the letter f and written as H(f). This is how it is determined:

$$H(f) = (|X(f)|^2) / (|X(f)|^2 + |N(f)|^2 / |F|^2) \quad (1)$$

Where $|X(f)|^2$ and $|N(f)|^2$ represent the power spectrum of the original image, on the one hand, and the noise, on the other, and $|F|^2$ indicates the strength of the image after it has been transformed using the Fast Fourier Method.

The filtered image in the frequency domain, denoted by $X_w(f)$, is obtained by multiplying the Wiener filter transfer function with the fast Fourier transform of the noisy image:

$$X_w(f) = H(f) * Y(f) \tag{2}$$

Finally, the filtered image in the spatial domain, denoted by x_w, is attained by applying the inverse Fast Fourier transform to $X_w(f)$:

$$x_w = Inverse\ Fast\ Fourier\ Transform\ (X_w(f)) \tag{3}$$

The resulting x_w represents the estimated original image after applying the application of the Wiener filter on the distorted image.

3.1.1. Wiener Filter in the Fast Fourier Domain

The Wiener filter is a linear filter, so it can be implemented using a convolution operation. The convolution operation can be executed using Fast Fourier Transform (FFT) algorithms, which can make it computationally effective for large images.

$$H(\omega) = K * |S(\omega)|^2 / |N(\omega)|^2 \tag{4}$$

Where

H(ω) is the Fourier transform of the Wiener filter coefficients.

K is a constant that depends on the desired signal-to-noise ratio (SNR).

S(ω) is an image transformation using a Fast Fourier transform.

N(ω) is An analysis of the noise using the Fast Fourier transform.

Since it achieves the best results, the Wiener filter is the preferred option in terms of minimising the MSE across the filtered image and the original image. However, it can be computationally expensive to implement, especially for large images. It is often used in conjunction with other techniques, such as non-linear filters, to improve the overall performance.

3.2. Multi-Resolution Convolutional Neural Network

An MRCNN is a kind of deep learning architecture that uses multiple Convolutional Neural Networks (CNNs) to process data at different resolutions. It incorporates multiple levels of resolutions in its feature extraction process. It is designed to handle images or data with varying levels of details or scales effectively. MRCNN is to capture both global and local information present in an input image. By processing the input at multiple resolutions or scales, the network can simultaneously consider different levels of details, allowing it to detect objects or patterns at various scales within the image.

3.2.1. Input Processing

The input image is typically passed through an initial set of convolutional and pooling layers to extract basic features. These layers are usually shared across all resolutions.

3.2.2. Multi-Resolution Branches

The network branches out into multiple pathways, each handling a different resolution of the input image. These branches are created by either downsampling or upsampling the feature maps from the previous layer. Downsampling reduces the resolution, while upsampling increases it, as shown in Figure 2.

3.2.3. Feature Extraction

Each resolution branch performs convolutional and pooling operations specific to its resolution level. The idea is to capture features at different scales or levels of abstraction. The branches can have varying depths, allowing for more complex feature extraction as the network goes deeper.

3.2.4. Fusion of Features

After feature extraction, the branches' outputs are combined or fused to aggregate the information learned at different resolutions. This fusion can happen through concatenation, element-wise summation, or other fusion techniques.

3.2.5. Classification or Regression

The fused features are then fed into fully connected layers or other output layers for the final classification or regression task. These layers perform task-specific computations, such as predicting object classes or estimating numerical values, as shown in Fig 3.

By considering multiple resolutions, a Multi-Resolution CNN can capture both fine-grained details and global context simultaneously. This makes it well-suited for activities like object identification and semantic segmentation, among others and image classification in scenarios where objects or patterns can appear at different scales or sizes.

The MR-CNN combines the use of multiple resolutions to capture and analyze different levels of information in an image. Here are the main mathematical equations involved in the MR-CNN:

Pseudocode for Wiener filter in the FFT

```
def wiener_filter(image, noise):
        S = np.fft.fft2(image)
        N = np.fft.fft2(noise)
        K = 1 / np.mean(N**2)
        H = K * S**2 / N**2
        h = np.fft.ifft2(H)
        return h
```

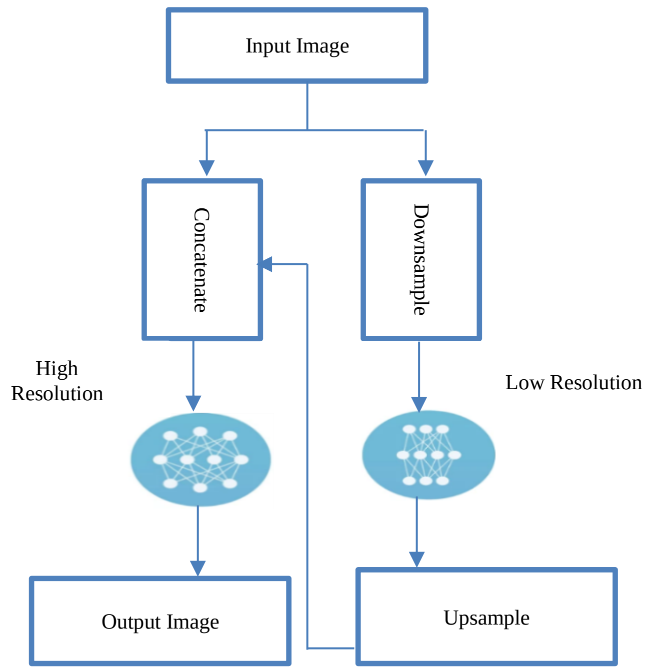

Fig. 2 MR-CNN Model

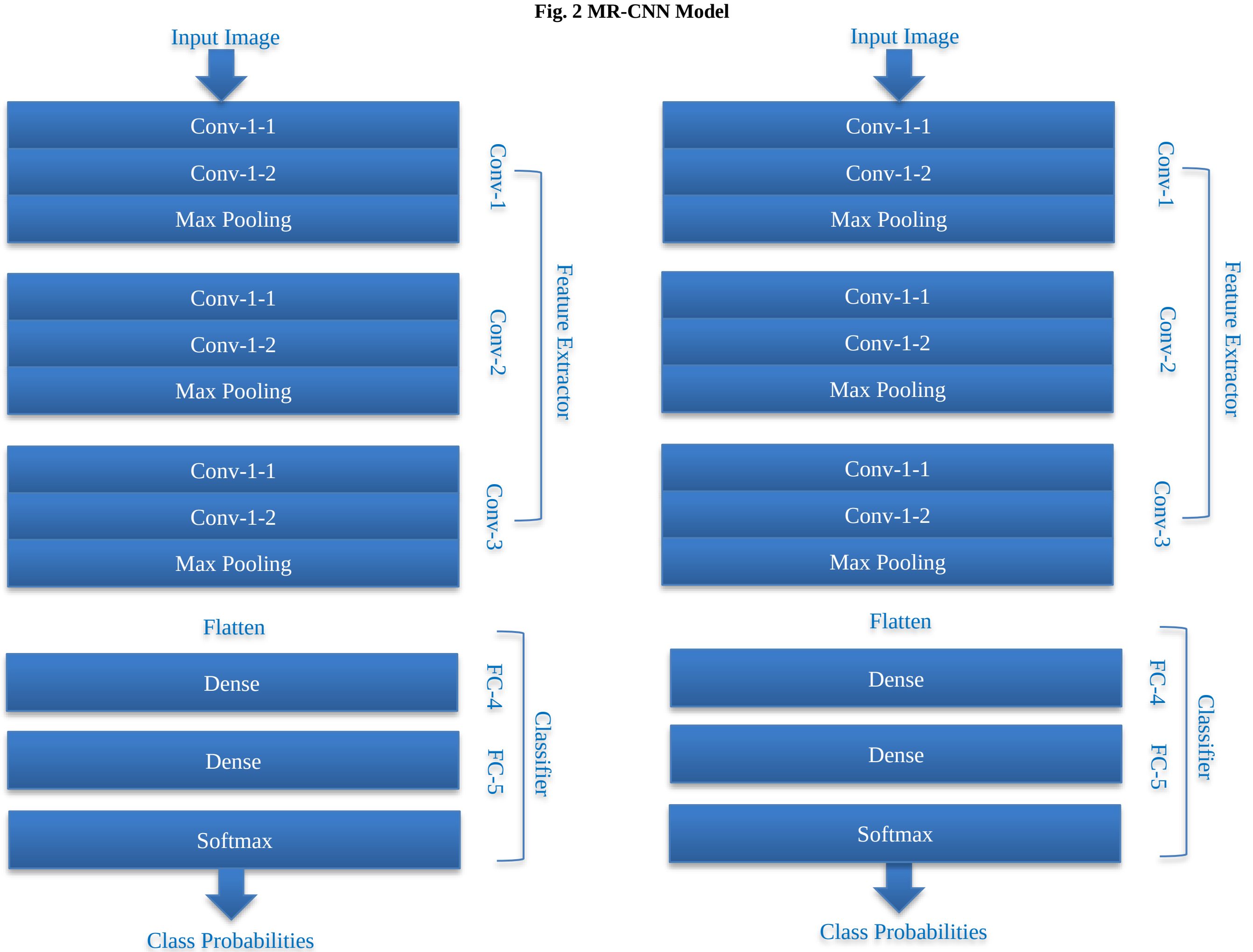

Fig. 3 MR-CNN Structure with different resolutions (480x640) and (1024x1280)

Convolutional Layer

In a typical convolutional layer of an MR-CNN, the output feature maps are generated by convolving the input feature maps with a set of learnable filters. The mathematical equation for the convolution operation can be represented as follows:

$$Y(i,j,k) = \sum[m,n,l]X(i+m,j+n,l) * W(m,n,l,k) + b(k) \quad (5)$$

Where Y(i, j, k) represents the position of the output feature (i, j) and, channel k, X(i+m, j+n, l) represents the input feature map at position (i+m, j+n), and channel l, W(m, n, l, k) represents the convolutional kernel/filter weights, and b(k) represents the bias term for channel k.

Pooling Layer

The pooling operation is commonly used in CNNs to lower the spatial dimensions while also downsampling the feature maps. Max pooling is a popular choice, and its mathematical equation can be written as follows:

$$Y(i,j,k) = max\big([m,n]X(si+m,sj+n,k)\big) \quad (6)$$

Where Y(i, j, k) represents the pooled output at position (i, j) and channel k, X(si+m, sj+n, k) represents the input feature map at position (si+m, sj+n) and channel k, and s is the pooling stride.

Non-linear Activation Function

Activation functions introduce the introduction of non-linearity into the network and the facilitation of the network's learning of complicated representations. Common choices include ReLU (Rectified Linear Unit) and its variants. The ReLU activation function can be defined as:

$$f(x) = max(0,x) \quad (7)$$

Where x represents the input to the activation function.

Multi-Resolution Integration

In an MR-CNN, the network typically includes branches or pathways that process inputs at different resolutions. These pathways can involve different numbers of layers and different filter sizes. The outputs from these pathways are then combined or concatenated at a later stage to form a unified feature representation. One way to mathematically represent this integration is by concatenating the feature maps from different pathways or by summing them element-wise. The exact formulation depends on the specific architecture and design choices of the MR-CNN.

The MR-CNN consists of multiple pathways, each processing inputs at a specific resolution. Each pathway can have its own set of convolutional layers, pooling layers, and other network components. The pathways capture different levels of information by operating on different scales or resolutions of the input data.

 # Pseudocode for a basic MR-CNN architecture

```
function convolution(input, weights, biases):
    output = empty array
    for each filter in weights:
        convolved = convolution_operation(input, filter)
        convolved = convolved + biases
        output.append(convolved)
    return output
function pooling(input, stride):
    output = empty array
    for each channel in input:
        pooled = max_pooling_operation(channel, stride)
        output.append(pooled)
    return output
function activation(input):
    output = relu(input)
    return output
function integrate(feature_maps): # Multi-Resolution Integration
    output = concatenate(feature_maps)  # Or element-wise sum, depending on the design
    return output
function MR_CNN(input): # MR-CNN Architecture
    # Pathway 1: High-resolution pathway
    conv1_1 = convolution(input, weights1_1, biases1_1)
    conv1_2 = convolution(conv1_1, weights1_2, biases1_2)
    pool1 = pooling(conv1_2, stride1)
    act1 = activation(pool1)
    # Pathway 2: Low-resolution pathway
    conv2_1 = convolution(input, weights2_1, biases2_1)
    pool2 = pooling(conv2_1, stride2)
    act2 = activation(pool2)
    # Integration of pathways
    integrated = integrate([act1, act2])
    # Additional layers for classification/regression
    ...
    return output
```

4. Results and Discussions

On MATLAB R2019b, a series of tests utilising the suggested approach for HMF-DWM denoising were carried out. The arrangement of the term "workstation" refers to the computer system with I(R) Xeon(R) CPU E5 1620 v4 @ 3.5 GHz, 64 GB RAM, and Windows 10.

4.1. Dataset Description

The dataset was collected from https://www.kaggle.com/datasets/slavkoprytula/aquarium-data-cots. The collection consists of seven different categories of marine life, and the positions of each individual animal's boxes are specified. Train, validate, and test sets have already been extracted from the dataset. It has 638 different images. There are a total of three cameras and three LED lights that are permanently attached as part of the camera configuration that was utilised to record the data.

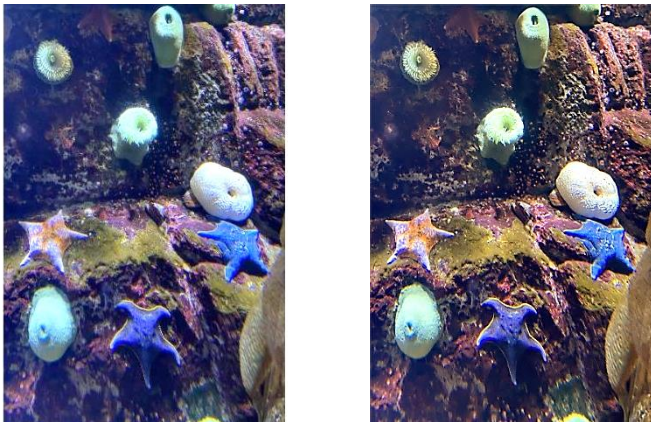

Fig. 4 (a) Input image (b) Pre-Processed image

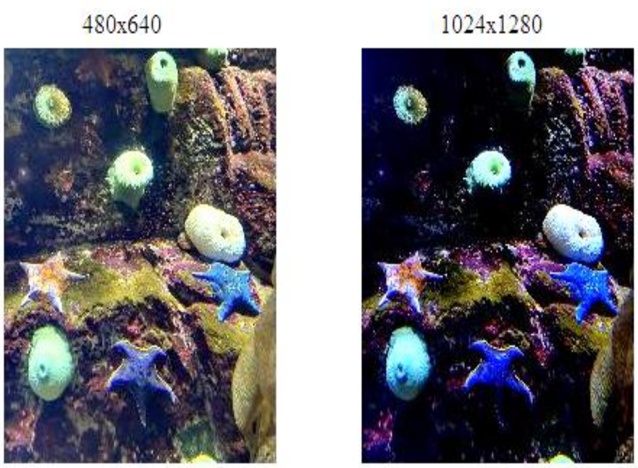

Fig. 5 Image with multi-resolution model

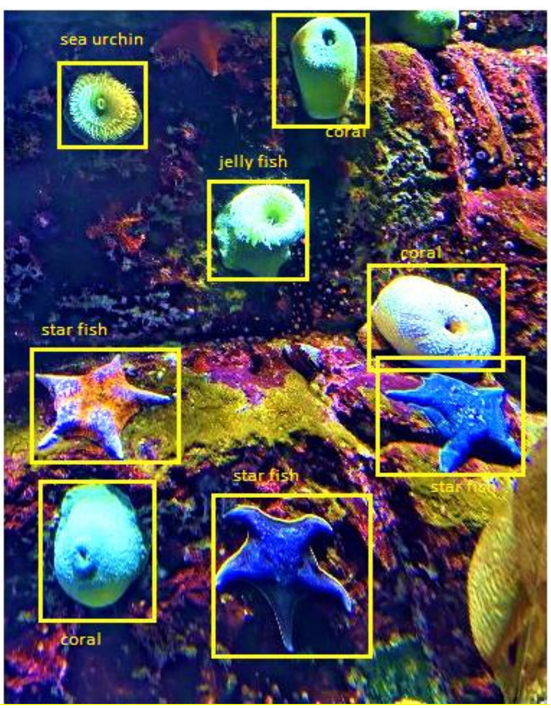

Fig. 6 Object detection using MR-CNN

4.2. Experimental Results

The simulated results are shown in Fig. 4-9. Fig 4(a) shows an input image underwater, and 4(b) shows a pre-processed image using the Weiner filter. Fig 5 presents the MR-CNN model with a resolution of the input image. Finally, Figure 6 detects an object underwater using our proposed algorithm.

4.3. Performance Analysis of Wiener Filter

Measurements that include the PSNR, the Structural Similarity Index (SSIM), and the RMSE are often used in order to assess the effectiveness of the Wiener filter. The quality of the image that has been recovered is evaluated using these criteria by contrasting it to the image that was captured before any noise was added. Higher values of PSNR or SSIM indicate better restoration quality. A lower RMSE value indicates that the Wiener filter has effectively reduced noise and improved the similarities that may be seen between the filtered and original images.

The formula for Peak Signal-to-Noise Ratio (PSNR) is as follows:

$$PSNR = 20 * log10(MAX / sqrt(MSE)) \quad (8)$$

Where,

The significance of the image pixel denoted by MAX is considered to be its highest conceivable value. The maximum value that may be represented in a grayscale image with 8 bits is 255.

The term "Mean Squared Error" (MSE) refers to the error that results from calculating the average of the squared disparities that exist between matching pixels in the original and reconstructed images.

The ratio of the greatest potential signal power (represented by MAX) to the power of the noise (represented by MSE) is what the PSNR number represents. Evaluation of the quality of reconstructed or processed images is a typical use of this technique, where higher PSNR values indicate better image quality and lower levels of distortion or noise.

The formula for RMSE in image processing is as follows:

$$RMSE = sqrt\left(\left(\frac{1}{N}\right) * \Sigma\left(I_{original} - I_{filtered}\right)^2\right) \quad (9)$$

Where

A metric that quantifies the degree to which the original and filtered images vary. The pixel value in the original image is represented by the variable I_original.

Table 1. PSNR and RMSE comparison

Filters	PSNR	RMSE
Median Filter	35.02	6.8
Gaussian Filter	34.29	6.4
Average Filter	34.29	6.1
Weiner Filter	35.26	5.89

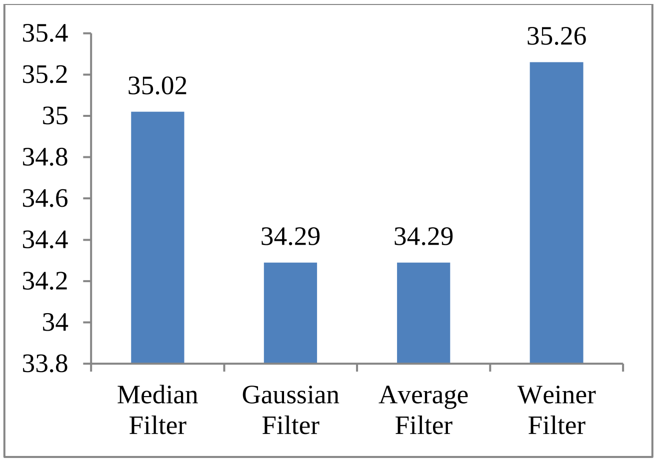

Fig. 7 Comparison of PSNR

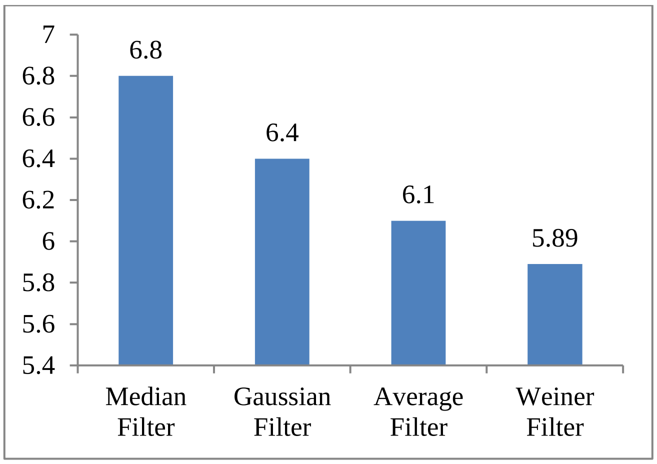

Fig. 8 Comparison of RMSE

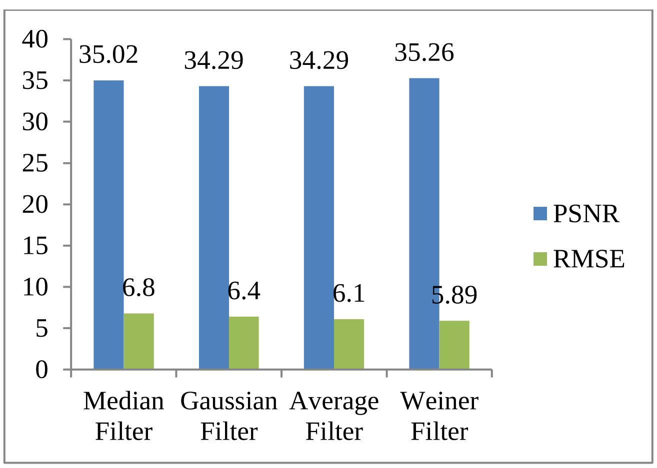

Fig. 9 Comparison of PSNR and RMSE

The value of the RMSE offers an estimate of the average difference between the unfiltered and filtered images. It provides a numerical representation of the degree to which the filtered image exhibits distortion or noise in comparison to the original image. A lower RMSE value indicates better similarity and closer resemblance between the two images, implying higher quality or accuracy in the filtering process.

Table 1 contains PSNR and RMSE for the Median filter, Gaussian filter and Average filter with Wiener filter, and it is graphically represented in Figures 7 and 8.

The Wiener filter can be influenced by various factors, including computational complexity, statistical characteristics of the signal, as well as the noise and the specific characteristics of the image and noise.

While it provides a powerful framework for image restoration, the efficiency of the Wiener filter should be considered in the context of the specific application requirements and constraints. Fig 9 displays a comparison of the PSNR and RMSE values for Gaussian Noise for each of the four distinct filters.

4.4. Performance Analysis of MR-CNN

The performance evaluation is shown in Fig 10, which clearly shows that our proposed model provides better accuracy than other models. Figure 11 depicts the mean average precision(mAP) for each iteration and compares mAP during training using a multi-resolution model against training without a multi-resolution model, which improved by 5%.

In summary, our approach surpasses existing techniques by intelligently integrating a tailored Hybrid Wiener Filter with a purpose-built MR-CNN, effectively addressing the limitations of conventional methods and deep learning approaches when applied individually to underwater object detection. The synergy between image enhancement and deep learning, coupled with strategies to mitigate data scarcity, forms the cornerstone of our methodology's success in achieving superior results reported in the literature.

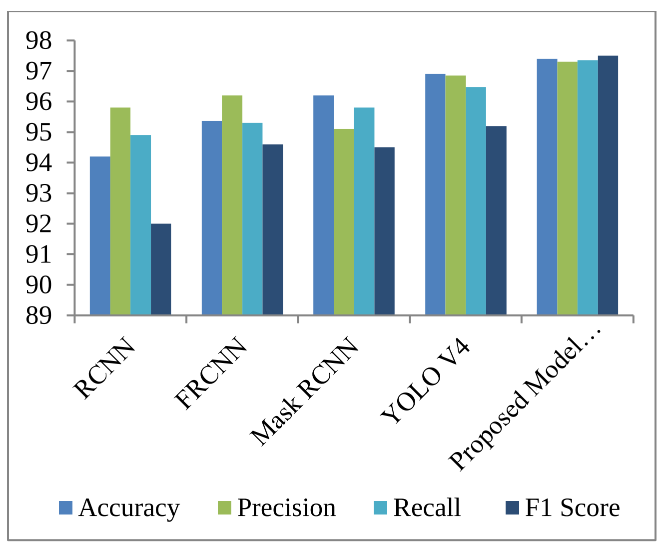

Fig. 10 Performance analysis

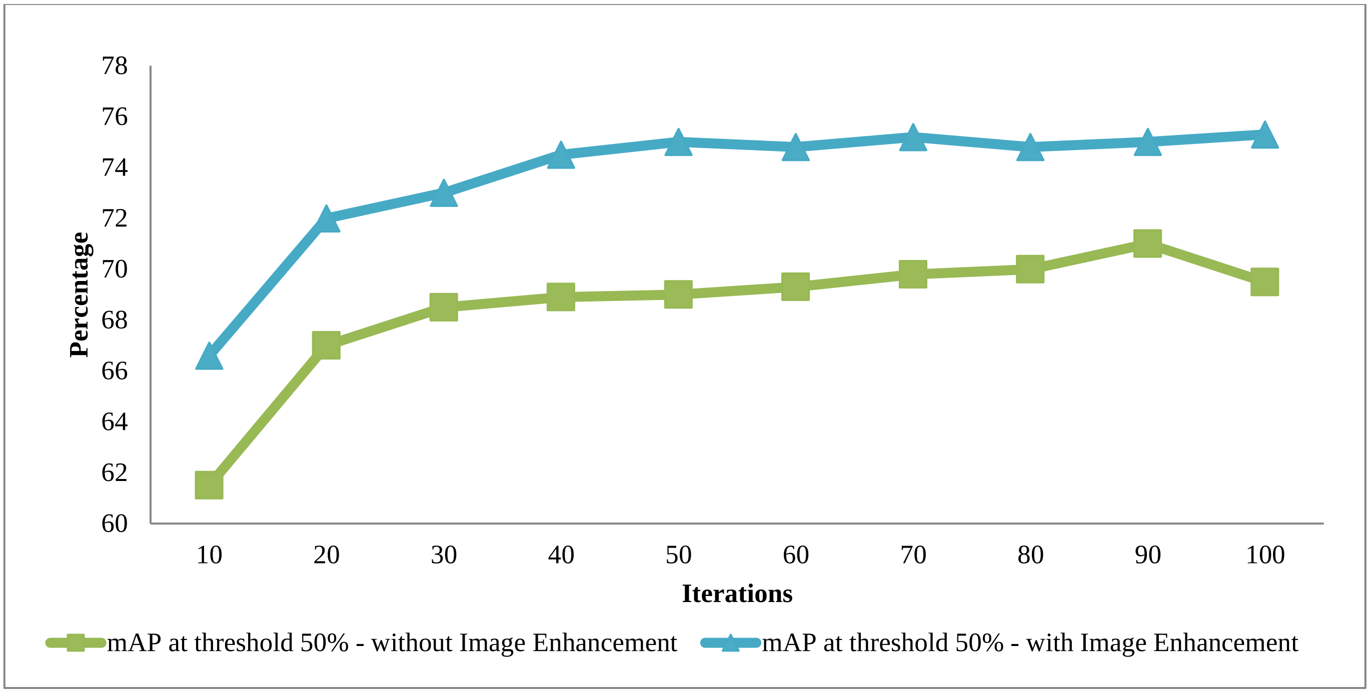

Fig. 11 Mean average precision

5. Conclusion

In this study, we suggested a unique method for the identification of objects in underwater environments using the MR-CNN, which integrates the Weiner filter with it. Image processing. The objective was to leverage the strengths of both classical image processing techniques and deep learning algorithms to improve object detection accuracy and robustness in challenging underwater environments. Experimental evaluations conducted on a diverse dataset of underwater images demonstrated the efficiency of the method that has been suggested. The effects showcased improved object detection accuracy and robustness compared to traditional methods and standalone deep learning approaches. The integration of the Weiner filter and MR-CNN contributed to better noise reduction, enhanced feature extraction, and more reliable object localization in challenging underwater conditions.

References

[1] Lintao Peng, Chunli Zhu, and Liheng Bian, "U-Shape Transformer for Underwater Image Enhancement," *IEEE Transactions on Image Processing*, vol. 32, pp. 3066-3079, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Jian Zhang et al., "An Improved YOLOv5-Based Underwater Object-Detection Framework," *Sensors*, vol. 23, no. 7, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Shubo Xu et al., "A Systematic Review and Analysis of Deep Learning-based Underwater Object Detection," *Neurocomputing*, vol. 527, pp. 204-232, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Boyang Sun et al., "UMGAN: Underwater Image Enhancement Network for Unpaired Image-to-Image Translation," *Journal of Marine Science and Engineering*, vol. 11, no. 2, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Prem Kumari Verma, Nagendra Pratap Singh, and Divakar Yadav, "Image Enhancement: A Review," *Ambient Communications and Computer Systems*, vol. 1097, pp. 347-355, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yang Wang et al., "A Deep CNN Method for Underwater Image Enhancement," *IEEE International Conference on Image Processing (ICIP)*, Beijing, China, pp. 1382-1386, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[7] Xiaobo Zhang, "Image Denoising Using Local Wiener Filter and its Method Noise," *Optik*, vol. 127, no. 17, pp. 6821-6828, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Fei Wu et al., "Adaptive Wiener Filter and Natural Noise to Eliminate Adversarial Perturbation," *Electronics*, vol. 9, no. 10, pp. 1-14, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Xin Sun, Lipeng Liu, and Junyu Dong, "Underwater Image Enhancement with Encoding-Decoding Deep CNN Networks," *IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computed, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, San Francisco, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] Meicheng Zheng, and Weilin Luo, "Underwater Image Enhancement Using Improved CNN Based Defogging," *Electronics*, vol. 11, no. 1, pp. 1-21, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Lanyong Zhang, Chengyu Li, and Hongfang Sun, "Object Detection/Tracking Toward Underwater Photographs by Remotely Operated Vehicles (ROVs)," *Future Generation Computer Systems*, vol. 126, pp. 163-168, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Radhwan Adnan Dakhil, and Ali Retha Hasoon Khayeat, "Review on Deep Learning Technique for Underwater Object Detection," *arXiv*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ganesh Babu Loganathan et al., "To Develop Multi-Object Detection and Recognition Using Improved GP-FRCNN Method," *8th International Conference on Smart Structures and Systems (ICSSS)*, Chennai, India, pp. 1-7. 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Keshetti Sreekala et al., "Deep Convolutional Neural Network with Kalman Filter Based Objected Tracking and Detection in Underwater Communications," *Wireless Networks*, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Maofa Wang et al., "Study on Underwater Target Tracking Technology Based on an LSTM–Kalman Filtering Method," *Applied Sciences*, vol. 12, no. 10, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Suja Cherukullapurath Mana, and T. Sasipraba, "An Intelligent Deep Learning Enabled Marine Fish Species Detection and Classification Model," *International Journal on Artificial Intelligence Tools*, vol. 31, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Joaquin Gabaldon et al., "Computer-Vision Object Tracking for Monitoring Bottlenose Dolphin Habitat Use and Kinematics," *Plos One*, vol. 17, no. 2, pp. 1-23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Atif Naseer et al., "A Novel Detection Refinement Technique for Accurate Identification of Nephrops Norvegicus Burrows in Underwater Imagery," *Sensors*, vol. 22, no. 12, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Rudrika Kalsotra, and Sakshi Arora, "Background Subtraction for Moving Object Detection: Explorations of Recent Developments and Challenges," *The Visual Computer*, vol. 38, no. 12, pp. 4151-4178, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Ajisha Mathias, Samiappan Dhanalakshmi, and R. Kumar, "Occlusion Aware Underwater Object Tracking Using Hybrid Adaptive Deep SORT-YOLOv3 Approach," *Multimedia Tools and Applications*, vol. 81, no. 30, pp. 44109-44121, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jiayi Fan et al., "Improvement of Object Detection Based on Faster R-CNN and YOLO," *36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, Jeju, Korea (South), pp. 1-4, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yu-Hsien Lin, Chao-Ming Yu, and Chia-Yu Wu, "Towards the Design and Implementation of an Image-Based Navigation System of an Autonomous Underwater Vehicle Combining a Color Recognition Technique and a Fuzzy Logic Controller," *Sensors*, vol. 21, no. 12, pp. 1-23, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Daoliang Li et al., "Automatic Counting Methods in Aquaculture: A Review," *Journal of the World Aquaculture Society*, vol. 52, no. 2, pp. 269-283, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Xianghui Li et al., "Intelligent Detection of Underwater Fish Speed Characteristics Based on Deep Learning," *5th Asian Conference on Artificial Intelligence Technology (ACAIT)*, Haikou, China, pp. 182-189, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Panpan Lu, Junyu Dong, and Xin Sun "Simple Online and Real-Time Underwater Multiple Object Tracking," *Twelfth International Conference on Graphics and Image Processing (ICGIP 2020)*, Xi'an, China, vol. 11720, pp. 107-111, 2021. [CrossRef] [Google Scholar] [Publisher Link]